TEAM 8

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/790,043*

DATE: 03/13/97   #2

TIME: 10:20:52

INPUT SET: S16152.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

ENTERED

```
 1                              SEQUENCE LISTING
 2
 3    (1)      General Information
 4
 5             (i) APPLICANT: Lonsdale, John
 6                            Milner, Peter
 7                            Payne, David
 8                            Pearson, Stewart
 9
10             (ii) TITLE OF THE INVENTION: Novel FabI
11
12
13             (iii) NUMBER OF SEQUENCES: 2
14
15             (iv) CORRESPONDENCE ADDRESS:
16                (A) ADDRESSEE: SmithKline Beecham Corporation
17                (B) STREET: 709 Swedeland Road
18                (C) CITY: King of Prussia
19                (D) STATE: PA
20                (E) COUNTRY: USA
21                (F) ZIP: 19406-0939
22
23             (v) COMPUTER READABLE FORM:
24                (A) MEDIUM TYPE: Diskette
25                (B) COMPUTER: IBM Compatible
26                (C) OPERATING SYSTEM: DOS
27                (D) SOFTWARE: FastSEQ for Windows Version 2.0
28
29             (vi) CURRENT APPLICATION DATA:
30                (A) APPLICATION NUMBER:
31                (B) FILING DATE: 24-JAN-1997
32                (C) CLASSIFICATION:
33
34             (vii) PRIOR APPLICATION DATA:
35                (A) APPLICATION NUMBER: 60/024845
36                (B) FILING DATE: 28-AUG-1996
37
38
39
40             (viii) ATTORNEY/AGENT INFORMATION:
41                (A) NAME: Gimmi, Edward R
42                (B) REGISTRATION NUMBER: 38,891
43                (C) REFERENCE/DOCKET NUMBER: GM50005
44
45             (ix) TELECOMMUNICATION INFORMATION:
46                (A) TELEPHONE: 610-270-4478
```

```
47              (B) TELEFAX: 610-270-5090
48              (C) TELEX:
49
50
51          (2) INFORMATION FOR SEQ ID NO:1:
52
53          (i) SEQUENCE CHARACTERISTICS:
54              (A) LENGTH: 256 amino acids
55              (B) TYPE: amino acid
56              (C) STRANDEDNESS: single
57              (D) TOPOLOGY: linear
58
59          (ii) MOLECULE TYPE: protein
60
61          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
62
63      Met Leu Asn Leu Glu Asn Lys Thr Tyr Val Ile Met Gly Ile Ala Asn
64       1               5                  10              15
65      Lys Arg Ser Ile Ala Phe Gly Val Ala Lys Val Leu Asp Gln Leu Gly
66                  20              25              30
67      Ala Lys Leu Val Phe Thr Tyr Arg Lys Glu Arg Ser Arg Lys Glu Leu
68              35              40              45
69      Glu Lys Leu Leu Glu Gln Leu Asn Gln Pro Glu Ala His Leu Tyr Gln
70          50              55              60
71      Ile Asp Val Gln Ser Asp Glu Glu Val Ile Asn Gly Phe Glu Gln Ile
72      65              70              75                  80
73      Gly Lys Asp Val Gly Asn Ile Asp Gly Val Tyr His Ser Ile Ala Phe
74                  85              90              95
75      Ala Asn Met Glu Asp Leu Arg Gly Arg Phe Ser Glu Thr Ser Arg Glu
76                  100             105             110
77      Gly Phe Leu Leu Ala Gln Asp Ile Ser Ser Tyr Ser Leu Thr Ile Val
78              115             120             125
79      Ala His Glu Ala Lys Lys Leu Met Pro Glu Gly Gly Ser Ile Val Ala
80          130             135             140
81      Thr Thr Tyr Leu Gly Gly Glu Phe Ala Val Gln Asn Tyr Asn Val Met
82      145             150             155             160
83      Gly Val Ala Lys Ala Ser Leu Glu Ala Asn Val Lys Tyr Leu Ala Leu
84                  165             170             175
85      Asp Leu Gly Pro Asp Asn Ile Arg Val Asn Ala Ile Ser Ala Gly Pro
86                  180             185             190
87      Ile Arg Thr Leu Ser Ala Lys Gly Val Gly Gly Phe Asn Thr Ile Leu
88              195             200             205
89      Lys Glu Ile Glu Glu Arg Ala Pro Leu Lys Arg Asn Val Asp Gln Val
90          210             215             220
91      Glu Val Gly Lys Thr Ala Ala Tyr Leu Leu Ser Asp Leu Ser Ser Gly
92      225             230             235             240
93      Val Thr Gly Glu Asn Ile His Val Asp Ser Gly Phe His Ala Ile Lys
94                  245             250             255
95
96          (2) INFORMATION FOR SEQ ID NO:2:
97
98          (i) SEQUENCE CHARACTERISTICS:
99              (A) LENGTH: 771 base pairs
```

INPUT SET: S16152.raw

```
100              (B) TYPE: nucleic acid
101              (C) STRANDEDNESS: double
102              (D) TOPOLOGY: linear
103
104          (ii) MOLECULE TYPE: Genomic DNA
105
106          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
107
108   ATGTTAAATC TTGAAAACAA AACATATGTC ATCATGGGAA TCGCTAATAA GCGTAGTATT       60
109   GCTTTTGGTG TCGCTAAAGT TTTAGATCAA TTAGGTGCTA AATTAGTATT TACTTACCGT      120
110   AAAGAACGTA GCCGTAAAGA GCTTGAAAAA TTATTAGAAC AATTAAATCA ACCAGAAGCG      180
111   CACTTATATC AAATTGATGT TCAAAGCGAT GAAGAGGTTA TTAATGGTTT TGAGCAAATT      240
112   GGTAAAGATG TTGGCAATAT TGATGGTGTA TATCATTCAA TCGCATTTGC TAATATGGAA      300
113   GACTTACGCG GACGCTTTTC TGAAACTTCA CGTGAAGGCT TCTTGTTAGC TCAAGACATT      360
114   AGTTCTTACT CATTAACAAT TGTGGCTCAT GAAGCTAAAA AATTAATGCC AGAAGGTGGT      420
115   AGCATTGTTG CAACAACATA TTTAGGTGGC GAATTCGCAG TTCAAAATTA TAATGTGATG      480
116   GGTGTTGCTA AAGCGAGCTT AGAAGCAAAT GTTAAATATT TAGCATTAGA CTTAGGTCCT      540
117   GATAATATTC GCGTTAATGC AATTTCAGCT GGTCCAATCC GTACATTAAG TGCAAAAGGT      600
118   GTGGGTGGTT TCAATACAAT TCTTAAAGAA ATCGAAGAGC GTGCACCTTT AAAACGTAAC      660
119   GTTGATCAAG TAGAAGTAGG TAAAACAGCG GCTTACTTRT TAAGTGACTT ATCAAGTGGC      720
120   GTTACAGGTG AAAATATTCA TGTAGATAGC GGATTCCACG CAATTAAATA A              771
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/790,043*

INPUT SET: S16152.raw

Line	Error	Original Text